

Tilburg University

A New Method for Deriving Robust and Globalized Robust Solutions of Uncertain Linear Conic Optimization Problems Having General Convex Uncertainty Sets

Gorissen, B.L.; Ben-Tal, A.; Blanc, J.P.C.; den Hertog, D.

Publication date:
2012

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Gorissen, B. L., Ben-Tal, A., Blanc, J. P. C., & den Hertog, D. (2012). *A New Method for Deriving Robust and Globalized Robust Solutions of Uncertain Linear Conic Optimization Problems Having General Convex Uncertainty Sets*. (CentER Discussion Paper; Vol. 2012-076). Operations research.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

No. 2012-076

**A NEW METHOD FOR DERIVING ROBUST AND
GLOBALIZED ROBUST SOLUTIONS OF UNCERTAIN
LINEAR CONIC OPTIMIZATION PROBLEMS
HAVING GENERAL CONVEX UNCERTAINTY SETS**

By

Bram L. Gorissen, Aharon Ben-Tal,
Hans Blanc, Dick den Hertog

19 September, 2012

ISSN 0924-7815

A new method for deriving robust and globalized robust solutions of uncertain linear conic optimization problems having general convex uncertainty sets

Bram L. Gorissen^{1,2}, Aharon Ben-Tal³, Hans Blanc¹, Dick den Hertog¹

Abstract

We propose a new way to derive tractable robust counterparts of a linear conic optimization problem by using the theory of Beck and Ben-Tal [2] on the duality between the robust (“pessimistic”) primal problem and its “optimistic” dual. First, we obtain a new *convex* reformulation of the dual problem of a robust linear conic program, and then show how to construct the primal robust solution from the dual optimal solution. Our result allows many new uncertainty regions to be considered. We give examples of tractable uncertainty regions that were previously intractable. The results are illustrated by solving a multi-item newsvendor problem. We also propose a new globalized robust counterpart that is more flexible, and is tractable for general convex uncertainty sets and any convex distance function.

keywords: robust optimization; general convex uncertainty regions; linear conic optimization

JEL classification: C61

1 Introduction

Robust Optimization (RO) is a paradigm for dealing with uncertain data in an optimization problem. Parts of RO originate from the seventies and eighties [21, 22, 23, 20, 15], but most of the existing theory and applications followed after new results in the late nineties [7, 13]. Extensive overviews of RO are given in [6, 8]. The basic idea of RO is that constraints have to hold for all parameter realizations in some given uncertainty region.

Currently, two tractable methods to solve an RO problem can be distinguished. Both methods are applied constraint-wise, i.e. they reformulate single constraints. The first method uses conic duality (e.g. used in [6]), while the second method uses Fenchel duality [4]. For some cases of uncertainty sets, both methods may not produce explicit tractable robust counterparts.

We present a new method for linear conic optimization problems, based on the result “primal worst equals dual best”. Our method gives tractable optimization problems for general convex uncertainty regions, and does not require the support function of the uncertainty region. We give examples of new uncertainty regions that

¹Department of Econometrics and Operations Research, Tilburg University.

²Corresponding author, Tilburg University, Department of Econometrics and Operations Research, 5000 LE Tilburg, the Netherlands. E-mail: b.l.gorissen@tilburguniversity.edu.

³Department of Industrial Engineering and Management, Technion - Israel Institute of Technology. CentER Extramural Fellow, Tilburg University. Research supported by the Israel–USA Science Foundation (BSF) Grant no. 2008302.

were previously intractable in Section 3. Our method also has advantages in both the formulation and in the complexity, which we treat in more detail in Section 2.

We also propose a new globalized robust counterpart (GRC) of a linear conic program (LCP) with two convex uncertainty regions, where the constraint holds for the smaller uncertainty region, and the violation of the constraint for the larger region is bounded by a convex distance function. This GRC is more flexible than the one in the literature. We show that our GRC can be formulated as an ordinary robust LCP with a (different) convex uncertainty region, which implies that it can be solved with the method presented in this paper.

For (LCP) with its associated dual (D-LCP), we will use the prefix D for the dual, O for the optimistic counterpart, and R for the robust counterpart. In (R-LCP), the constraints have to hold for all uncertain parameters in some given uncertainty region, whereas in (OD-LCP), the constraints have to hold for only a single value of each uncertain parameter. The result by Beck and Ben-Tal [2] is that (R-LCP) and (OD-LCP) are dual to each other. Moreover, there is no duality gap when (OD-LCP) is bounded and satisfies the Slater condition.

(OD-LCP) contains products of variables and is in general nonconvex. Beck and Ben-Tal [2] use ad-hoc reformulations tailored to specific uncertainty regions to obtain convex optimization problems. Our contribution is giving a tractable convex formulation of (OD-LCP) for any convex uncertainty region, and showing how to translate an optimal solution of (OD-LCP) to an optimal solution of (R-LCP). This enables us to solve robust counterparts of LCPs, for uncertainty sets that were previously out of reach.

Our method uses the dual problem and can therefore not directly be applied to problems with integer variables. However, our method can be used to solve LP relaxations. Therefore, existing cutting plane and branch & bound methods can be applied to the primal problem. Warm start strategies can still be applied.

2 Our method

Consider the following Linear Conic Program (LCP):

$$\begin{aligned}
 \text{(LCP)} \quad & \max_{\mathbf{x}} \quad \mathbf{c}^\top \mathbf{x} \\
 \text{s.t.} \quad & \mathbf{a}_i^\top \mathbf{x} \leq b_i \quad \forall i \\
 & \mathbf{x} \in \mathcal{K},
 \end{aligned}$$

where \mathcal{K} is a closed convex cone. If \mathcal{K} is the nonnegative orthant \mathbb{R}_+^n , (LCP) is an LP in canonical form. Other common choices for \mathcal{K} are the second-order cone and the semidefinite cone. The dual of (LCP) is given by:

$$\begin{aligned}
 \text{(D-LCP)} \quad & \min_{\mathbf{y}} \quad \mathbf{b}^\top \mathbf{y} \\
 \text{s.t.} \quad & \sum_i y_i \mathbf{a}_i - \mathbf{c} \in \mathcal{K}^* \\
 & \mathbf{y} \geq 0,
 \end{aligned}$$

where \mathcal{K}^* is the dual cone of \mathcal{K} . Assume that \mathbf{a}_i are uncertain, but known to reside in some convex compact uncertainty region $\mathcal{U}_i = \{\mathbf{a}_i : f_{ik}(\mathbf{a}_i) \leq 0 \quad \forall k\}$, where f_{ik} are given closed proper convex functions. Note that this description of \mathcal{U}_i also includes constraints of the form $D_i \mathbf{a}_i - \mathbf{d}_i \in \mathcal{C}_i$, where \mathcal{C}_i is a convex set, by using the indicator function $\delta(D_i \mathbf{a}_i - \mathbf{d}_i | \mathcal{C}_i)$, taking the value 0 if $D_i \mathbf{a}_i - \mathbf{d}_i$ is in \mathcal{C}_i , ∞ otherwise. The

robust counterpart of (D-LCP) is given by:

$$\begin{aligned}
(\text{R-LCP}) \quad & \max_{\mathbf{x}} \quad \mathbf{c}^\top \mathbf{x} \\
\text{s.t.} \quad & \mathbf{a}_i^\top \mathbf{x} \leq b_i \quad \forall \mathbf{a}_i : f_{ik}(\mathbf{a}_i) \leq 0 \quad \forall i \quad \forall k \\
& \mathbf{x} \in \mathcal{K},
\end{aligned}$$

whereas the optimistic counterpart of (D-LCP) is given by:

$$\begin{aligned}
(\text{OD-LCP}) \quad & \min_{\mathbf{a}_i, \mathbf{y}} \quad \mathbf{b}^\top \mathbf{y} \\
\text{s.t.} \quad & \sum_i y_i \mathbf{a}_i - \mathbf{c} \in \mathcal{K}^* \\
& f_{ik}(\mathbf{a}_i) \leq 0 \quad \forall i \quad \forall k \\
& \mathbf{y} \geq 0.
\end{aligned} \tag{1}$$

$$f_{ik}(\mathbf{a}_i) \leq 0 \quad \forall i \quad \forall k \tag{2}$$

$$\mathbf{y} \geq 0. \tag{3}$$

A result by Beck and Ben-Tal [2] is that (OD-LCP) is a dual problem of (R-LCP), and that if (OD-LCP) satisfies the Slater condition, the duality gap is 0. Less general but similar results can be found in [14, 19, 22, 23]. The values of (R-LCP) and (OD-LCP) are equal if (OD-LCP) is bounded and satisfies the Slater condition. For $\mathcal{K} = \mathbb{R}_+^n$, (OD-LCP) is called a Generalized LP (GLP) [12, p. 434]. It contains the product of variables $y_i \mathbf{a}_i$ and is in general nonconvex. The following lemma and its proof are based on [9, proposition 1].

Lemma 1 *The projection of the feasible region of (OD-LCP) on \mathbf{y} is convex, while the projection on \mathbf{a}_i is not necessarily convex.*

Proof. Let $(\{\mathbf{a}_i^1\}_i, \mathbf{y}^1)$ and $(\{\mathbf{a}_i^2\}_i, \mathbf{y}^2)$ be feasible for (OD-LCP), and let $\lambda \in [0, 1]$. Define $\mu_i = \lambda y_i^1 / (\lambda y_i^1 + (1 - \lambda) y_i^2)$ when $y_i^1 \neq 0$ or $y_i^2 \neq 0$, $\mu_i = 1$ otherwise. We claim that $(\{\mu_i \mathbf{a}_i^1 + (1 - \mu_i) \mathbf{a}_i^2\}_i, \lambda \mathbf{y}^1 + (1 - \lambda) \mathbf{y}^2)$ is feasible for (OD-LCP). It is obvious that (2) and (3) are satisfied. It remains to verify (1):

$$\begin{aligned}
& \sum_i (\lambda y_i^1 + (1 - \lambda) y_i^2) (\mu_i \mathbf{a}_i^1 + (1 - \mu_i) \mathbf{a}_i^2) - \mathbf{c} \\
&= \sum_i (\lambda y_i^1 \mathbf{a}_i^1 + (1 - \lambda) y_i^2 \mathbf{a}_i^2) - \mathbf{c} \\
&= \lambda \left(\sum_i y_i^1 \mathbf{a}_i^1 - \mathbf{c} \right) + (1 - \lambda) \left(\sum_i y_i^2 \mathbf{a}_i^2 - \mathbf{c} \right) \in \mathcal{K}^*,
\end{aligned}$$

proving convexity in \mathbf{y} . We now provide a counterexample for convexity in \mathbf{a}_i . Let (OD-LCP) be constrained by $y_1 a_1 + y_2 a_2 - 1 \in \mathbb{R}_+$, $\|a_i\|_2 \leq 5$, $\mathbf{y} \geq 0$. Both $(a_1, a_2) = (1, -1)$ and $(a_1, a_2) = (-1, 1)$ are feasible (with $\mathbf{y} = (1, 0)$ and $\mathbf{y} = (0, 1)$, respectively), but $(a_1, a_2) = (0, 0)$ is not. ■

Dantzig solves GLPs where the coefficients are restricted to linear functions f_{ik} with a decomposition method that has the advantage that both the master and the subproblems are LP [12, p. 435–437]. In the context of RO, a method that is likely to be faster and also results in LP problems has been developed in [23]. We do not consider these approaches because our result is more efficient.

Dantzig mentions substituting $\mathbf{v}_i = y_i \mathbf{a}_i$ and multiplying constraint (2) with y_i as a solution approach to GLPs [12, p. 434], which has already been applied to the dual of LPs with polyhedral uncertainty [19]. We will show that this approach is also valid for our problem and results in a convex optimization problem where the vectors \mathbf{a}_i do no longer appear. Dantzig notes that the resulting problem is only equivalent when $\mathbf{v}_i \neq \mathbf{0}$ is not possible if $y_i = 0$. In Lemma 2 we show that conic representable uncertainty regions satisfy this criterion. For another uncertainty region it may be necessary to take the intersection with an enclosing box, thereby adding linear constraints to

its description. This always results in an equivalent uncertainty region because the original uncertainty region is compact. After substituting $\mathbf{v}_i = y_i \mathbf{a}_i$ and multiplying constraint (2) with y_i , we get the following convex optimization problem, equivalent to (OD-LCP):

$$\begin{aligned} (\text{COD-LCP}) \quad & \min_{\mathbf{v}_i, \mathbf{y}} \quad \mathbf{b}^\top \mathbf{y} \\ & \text{s.t.} \quad \sum_i \mathbf{v}_i - \mathbf{c} \in \mathcal{K}^* \end{aligned} \quad (4)$$

$$\begin{aligned} & y_i f_{ik} \left(\frac{\mathbf{v}_i}{y_i} \right) \leq 0 \quad \forall i \quad \forall k \\ & \mathbf{y} \geq \mathbf{0}, \end{aligned} \quad (5)$$

where $0 f_{ik} \left(\frac{\mathbf{v}_i}{0} \right) = \lim_{y_i \downarrow 0} y_i f_{ik} \left(\frac{\mathbf{v}_i}{y_i} \right)$. (COD-LCP) is indeed a convex problem, since the perspective function $g_{ik}(\mathbf{v}_i, y_i) := y_i f_{ik}(\mathbf{v}_i/y_i)$ is convex on $\mathbb{R}^n \times \mathbb{R}_+$ [11]. We give a short proof for convexity on $\mathbb{R}^n \times \mathbb{R}_+ \setminus \{0\}$ that uses convex analysis:

$$\begin{aligned} & y_i f_{ik} \left(\frac{\mathbf{v}_i}{y_i} \right) \\ &= y_i f_{ik}^{**} \left(\frac{\mathbf{v}_i}{y_i} \right) \\ &= y_i \sup_{\mathbf{x}} \left\{ \frac{\mathbf{v}_i^\top}{y_i} \mathbf{x} - f_{ik}^*(\mathbf{x}) \right\} \\ &= \sup_{\mathbf{x}} \left\{ \mathbf{v}_i^\top \mathbf{x} - y_i f_{ik}^*(\mathbf{x}) \right\}, \end{aligned}$$

from which it follows that g_{ik} is jointly convex because it is the pointwise supremum of functions that are linear in \mathbf{v}_i and y_i .

While (R-LCP) is difficult to solve because it has an infinite number of constraints, (COD-LCP) does not have “for all” constraints. For some popular choices of f_{ik} for which an exact reformulation of (R-LCP) is known, (COD-LCP) is not more difficult to solve than (R-LCP). For instance, when the uncertainty region is polyhedral, (R-LCP) can be reformulated as an LP, and (COD-LCP) is also an LP. When the uncertainty region is an ellipsoid, (R-LCP) can be reformulated as a conic quadratic program, and (COD-LCP) is also a conic quadratic program. In general, the formulation of the perspective in (COD-LCP) for conic representable uncertainty regions is given by the following lemma:

Lemma 2 *Suppose the (bounded) uncertainty region \mathcal{U}_i has a conic representation: $\mathcal{U}_i = \{\mathbf{a}_i : D_i \mathbf{a}_i - \mathbf{d}_i \in \mathcal{K}_i\}$, where \mathcal{K}_i is a cone. Then the equivalent constraint (5) in (COD-LCP) is: $D_i \mathbf{v}_i - y_i \mathbf{d}_i \in \mathcal{K}_i$.*

Proof. (COD-LCP) has the constraint $y_i \delta(D_i \mathbf{v}_i/y_i - \mathbf{d}_i | \mathcal{K}_i) \leq 0$. For $y_i \geq 0$, this is equivalent to: $\delta(D_i \mathbf{v}_i - y_i \mathbf{d}_i | \mathcal{K}_i) \leq 0$, since \mathcal{K}_i is a cone. This can be formulated as $D_i \mathbf{v}_i - y_i \mathbf{d}_i \in \mathcal{K}_i$. It remains to verify that $y_i = 0$ implies $\mathbf{v}_i = \mathbf{0}$. If $\mathbf{v}_i^* \neq \mathbf{0}$ satisfies $D_i \mathbf{v}_i^* \in \mathcal{K}_i$, the uncertainty region recedes in the direction of \mathbf{v}_i^* , contradicting boundedness. ■

Corollary 1 *A conic quadratic representable uncertainty region results in conic quadratic constraints (5) in (COD-LCP).*

In practice it is often necessary to have the primal robust solution \mathbf{x} of (R-LCP), instead of a solution to (COD-LCP). The following theorem shows how \mathbf{x} can be recovered from an optimal solution of (COD-LCP).

Theorem 1 *Assume that (COD-LCP) is bounded and satisfies the Slater condition. A KKT vector of constraint (4) corresponds to an optimal solution \mathbf{x} of (R-LCP).*

Proof. First, we show that the dual variables associated with constraint (4) are the optimization variables of (R-LCP). The Lagrangian of (COD-LCP) is given by:

$$L(\mathbf{y}, v, \mathbf{x}, z) = \mathbf{b}^\top \mathbf{y} + \mathbf{x}^\top \left(\mathbf{c} - \sum_i \mathbf{v}_i \right) + \sum_{i,k} z_{ik} y_i f_{ik} \left(\frac{\mathbf{v}_i}{y_i} \right),$$

and hence, (R-LCP) is given by:

$$\begin{aligned} & \max_{\mathbf{x} \in \mathcal{K}, z \geq 0} \min_{\mathbf{y} \geq 0, v} L(\mathbf{y}, v, \mathbf{x}, z) \\ &= \max_{\mathbf{x} \in \mathcal{K}} \left\{ \mathbf{c}^\top \mathbf{x} + \max_{z \geq 0} \min_{\mathbf{y} \geq 0, v} \left\{ \mathbf{b}^\top \mathbf{y} - \sum_i \mathbf{v}_i^\top \mathbf{x} + \sum_{i,k} z_{ik} y_i f_{ik} \left(\frac{\mathbf{v}_i}{y_i} \right) \right\} \right\} \\ &= \max_{\mathbf{x} \in \mathcal{K}} \left\{ \mathbf{c}^\top \mathbf{x} + \max_{z \geq 0} \min_{\mathbf{y} \geq 0, a} \left\{ \sum_i y_i \left(b_i - \mathbf{a}_i^\top \mathbf{x} + \sum_k z_{ik} f_{ik}(\mathbf{a}_i) \right) \right\} \right\} \\ &= \max_{\mathbf{x} \in \mathcal{K}} \left\{ \mathbf{c}^\top \mathbf{x} : \max_{z \geq 0} \min_a \left\{ b_i - \mathbf{a}_i^\top \mathbf{x} + \sum_k z_{ik} f_{ik}(\mathbf{a}_i) \right\} \geq 0 \quad \forall i \right\} \\ &= \max_{\mathbf{x} \in \mathcal{K}} \left\{ \mathbf{c}^\top \mathbf{x} : \min_a \max_{z \geq 0} \left\{ b_i - \mathbf{a}_i^\top \mathbf{x} + \sum_k z_{ik} f_{ik}(\mathbf{a}_i) \right\} \geq 0 \quad \forall i \right\} \\ &= \max_{\mathbf{x} \in \mathcal{K}} \left\{ \mathbf{c}^\top \mathbf{x} : \mathbf{a}_i^\top \mathbf{x} \leq b_i \quad \forall \mathbf{a}_i : f_{ik}(\mathbf{a}_i) \leq 0 \quad \forall i \quad \forall k \right\}, \end{aligned}$$

where in the second equality the substitution $\mathbf{a}_i = \mathbf{v}_i/y_i$ is made, and the fourth equality is based on a min-max result for convex-concave functions [18, Corollary 37.3.2]. The problem in the last equality is indeed (R-LCP).

An optimal \mathbf{x} of (R-LCP) is therefore equal to the part of a KKT vector that relates to (4) [18, Corollary 28.4.1]. ■

This theorem is useful in practice because many solvers can output a KKT vector. There is also another way to obtain a solution of (R-LCP), similar to the method mentioned in [22]. The idea is to use the “dual best \mathbf{a}_i ” as the “primal worst \mathbf{a}_i ”: translate a solution of (COD-LCP) to a solution of (OD-LCP), then fix the variables \mathbf{a}_i , remove the constraints on \mathbf{a}_i , and dualize that problem with respect to y_i . The result is a problem similar to (LCP), where the vectors \mathbf{a}_i have been replaced with “worst case” \mathbf{a}_i . We call this problem (M-LCP). This method only works if (COD-LCP) has a unique optimal \mathbf{a}_i , and if (M-LCP) has a unique optimal \mathbf{x} [22]. Then, the value of (M-LCP) equals the value of (OD-LCP), and \mathbf{x} is both feasible and optimal for (R-LCP).

Our method has the following advantages:

1. We use the perspectives of the functions that define the uncertainty region, which are easy to formulate. Existing methods in RO use the conjugate of the perspective to reformulate the “for all” constraints in (R-LCP) [4], which may not result in closed-form formulations for many uncertainty regions. We give examples of new tractable uncertainty regions in Section 3.
2. The perspectives of the functions that define the uncertainty region are part of the optimization problem. This simplifies the derivation, especially when the uncertainty region is the intersection of several sets, which we show for the non-empty intersection of a polyhedron $D_i \mathbf{a}_i \leq \mathbf{d}_i$ and an ellipsoid $\|\mathbf{a}_i - \mathbf{a}_i^0\|_2 \leq \Omega$.

We obtain (R-LCP) after a tedious derivation using existing RO techniques:

$$\begin{aligned}
(\text{R-LCP}) \quad & \max_{\mathbf{s}, \mathbf{x}, \mathbf{y}} \quad \mathbf{c}^\top \mathbf{x} \\
\text{s.t.} \quad & (\mathbf{a}_i^0)^\top \mathbf{s}_i + \Omega \|\mathbf{s}_i\|_2 + \mathbf{d}_i^\top \mathbf{y}_i \leq b_i \quad \forall i \\
& D_i^\top \mathbf{y}_i + \mathbf{s}_i = \mathbf{x} \quad \forall i \\
& \mathbf{x} \in \mathcal{K}, \mathbf{y} \geq 0,
\end{aligned}$$

whereas with our method the problem simply has one constraint per uncertainty region:

$$\begin{aligned}
(\text{COD-LCP}) \quad & \min_{\mathbf{v}, \mathbf{y}} \quad \mathbf{b}^\top \mathbf{y} \\
\text{s.t.} \quad & \sum_i \mathbf{v}_i - \mathbf{c} \in \mathcal{K}^* \\
& D_i \mathbf{v}_i \leq y_i \mathbf{d}_i \quad \forall i \\
& \|\mathbf{v}_i - y_i \mathbf{v}_i^0\|_2 \leq \Omega y_i \quad \forall i \\
& \mathbf{y} \geq 0.
\end{aligned}$$

An additional advantage is that the formulation of (COD-LCP) directly reveals which uncertainty region was used.

3. There may be a computational advantage in the number of variables and constraints in (COD-LCP) compared to results obtained with existing RO techniques. The latter may e.g. require an explicit conic representation (e.g. see [6, Thm 1.3.4]) which can significantly increase the number of variables and constraints.
4. Another computational advantage arises for an ellipsoidal uncertainty region: $\{\mathbf{a}_i : \|Q(\mathbf{a}_i - \mathbf{a}_i^0)\|_2 \leq \Omega\}$. Current RO reformulates this to: $\{Q^{-1}\boldsymbol{\zeta} + \mathbf{a}_i^0 : \|\boldsymbol{\zeta}\|_2 \leq \Omega\}$. Our method does not require this reformulation, which has several advantages. First, computing Q^{-1} may be computationally challenging, e.g. if the dimensions of Q are large or if Q has a large condition number. Second, Q^{-1} may be dense while Q is sparse.

3 New tractable uncertainty regions

In this section we present examples of uncertainty regions for which the robust counterpart could previously not be obtained explicitly, but are now tractable.

1. The first example is given by problems in which several scenarios for the parameters can be distinguished, but the probabilities on these scenarios are not known. Suppose these unknown probabilities can be estimated based on historical data, and an optimization problem has a constraint involving these probabilities. An example of this is a constraint on expected value. For such problems, a wide class of uncertainty regions is given in terms of the distance between the real probability vector \mathbf{p} and an historical estimate $\hat{\mathbf{p}}$, both indexed by the scenario s :

$$\mathcal{U}_i = \left\{ \mathbf{p} : \mathbf{p} \geq \mathbf{0}, \sum_s p^{(s)} = 1, d(\mathbf{p}, \hat{\mathbf{p}}) \leq \rho \right\}, \quad (6)$$

where d is the distance measure and ρ is the level of uncertainty. Note that the constraint $\sum_s p^{(s)} = 1$ is not necessary for the following results to hold, so \mathbf{p} does not need to be a probability vector. We consider several classes of distance measures. The first class of distance measures that contains previously intractable cases is ϕ -divergence, which for a convex function ϕ that satisfies

$\phi(1) = 0$ is defined by $d(\mathbf{p}, \hat{\mathbf{p}}) = \sum_s \hat{p}^{(s)} \phi(p^{(s)}/\hat{p}^{(s)})$. Ben-Tal et al. [5] show how to choose ρ in (6) based on historical observations, and give tractable robust counterparts for several choices of ϕ . One example for which their method does not give a tractable reformulation is the Matusita distance [16], where $\phi(t) = |1 - t^\alpha|^{1/\alpha}$ for given $\alpha \in (0, 1)$.

The second class of distance measures is based on the Bregman distance [10], which is given by:

$$d(\mathbf{p}, \hat{\mathbf{p}}) = g(\mathbf{p}) - g(\hat{\mathbf{p}}) - (\nabla g(\hat{\mathbf{p}}))^\top (\mathbf{p} - \hat{\mathbf{p}}),$$

where g is real-valued, continuously-differentiable and strictly convex on the set of probability vectors. The Bregman distance is convex in its first argument. Previously, uncertainty regions were intractable for many choices of g , while with our results any g gives a tractable optimistic counterpart.

The third class of distance measures is the Rényi divergence [17]:

$$d(\mathbf{p}, \hat{\mathbf{p}}) = \frac{1}{\alpha - 1} \log \sum_i \left(\hat{p}_i^{(s)}\right)^\alpha \left(p_i^{(s)}\right)^{1-\alpha},$$

where $\alpha > 0$ and $\alpha \neq 1$. After some rewriting, an uncertainty region based on this distance measure can also be reformulated using Fenchel duality [4]. However, the rewriting is not always possible, e.g. when this divergence measure is clustered with other distance measures [1], while our result can then still be applied.

2. The second example of new tractable uncertainty regions is when the uncertainty region contains products of parameters:

$$\mathcal{U}_i = \left\{ (\mathbf{a}_i, \boldsymbol{\zeta}_i) : \boldsymbol{\zeta}_i \geq \mathbf{0}, \mathbf{a}_i = \mathbf{a}_i^0 + \sum_j \zeta_{ij} \mathbf{B}_{ij}, g_{ijk}(\mathbf{B}_{ij}) \leq 0, h_{ik}(\boldsymbol{\zeta}_i) \leq 0 \forall j, k \right\},$$

where g_{ijk} and h_{ik} are convex functions and $\mathbf{B}_{ij}, \mathbf{a}_i^0$ are vectors. The same substitution we applied to (OD-CLP), can also be applied to this uncertainty region. Let $\mathbf{v}_{ij} = \zeta_{ij} \mathbf{B}_{ij}$. The uncertainty region \mathcal{U}_i can be rewritten as:

$$\mathcal{U}_i = \left\{ (\mathbf{a}_i, \boldsymbol{\zeta}_i) : \boldsymbol{\zeta}_i > \mathbf{0}, \mathbf{a}_i = \mathbf{a}_i^0 + \sum_j \mathbf{v}_{ij}, \zeta_{ij} g_{ijk}(\mathbf{v}_{ij}/\zeta_{ij}) \leq 0, h_{ik}(\boldsymbol{\zeta}_i) \leq 0 \forall j, k \right\},$$

which is convex, and hence, leads to a tractable optimistic counterpart. We mention three cases where this uncertainty region appears. First, it appears in factor models with uncertainty in both $\boldsymbol{\zeta}_i$ and the model coefficients. Second, it appears in a constraint containing the steady-state distributions of a Markov chain, where the transition probabilities are uncertain. The uncertainty region then looks as follows:

$$\mathcal{U}_i = \left\{ \boldsymbol{\pi} \in \mathbb{R}_+^n : \mathbf{e}^\top \boldsymbol{\pi} = 1, \sum_j \pi_j \mathbf{B}_j = \boldsymbol{\pi}, g_{jk}(\mathbf{B}_j) \leq 0 \forall j, k \right\},$$

where \mathbf{B}_j are the columns of the matrix with transition probabilities. Markov chains with column-wise uncertainty in the transition matrix were also considered in [9]. Third, it appears in a constraint on the next time period probability vector \mathbf{p}_i of a Markov chain when there is uncertainty both in the transition matrix and in the current state:

$$\mathcal{U}_i = \left\{ \mathbf{p}_i \in \mathbb{R}_+^n : \mathbf{p}_i = \sum_j (\mathbf{p}_i^0)_j \mathbf{B}_j, g_{jk}(\mathbf{B}_j) \leq 0, h_k(\mathbf{p}_i^0) \leq 0 \forall j, k \right\},$$

where \mathbf{B}_j are the columns of the transition matrix and \mathbf{p}_i^0 is the current probability vector.

3. The third example of a new “more tractable” uncertainty region is a nonconvex region. It is well known that an equivalent constraint is obtained by replacing the uncertainty region with its convex hull (e.g. see [6, p. 12]). For example, $\mathcal{U}_i = \{\mathbf{a}_i : \mathbf{a}_i^\top Q_i \mathbf{a}_i \leq \Omega_i\}$, where $Q_i \in \mathbb{R}^{L \times L}$ is not positive semidefinite, may be replaced with:

$$\mathcal{U}_i = \left\{ \sum_{k=1}^{L+1} \lambda_i^{(k)} \mathbf{a}_i^{(k)} : \lambda_i \geq 0, \sum_{k=1}^{L+1} \lambda_i^{(k)} = 1, \left(\mathbf{a}_i^{(k)} \right)^\top Q_i \mathbf{a}_i^{(k)} \leq \Omega_i \forall k \right\}.$$

The part of the nonconvexity from $\lambda_i^{(k)} \mathbf{a}_i^{(k)}$ can be eliminated by substituting $w_{ik} = \lambda_i^{(k)} \mathbf{a}_i^{(k)}$:

$$\mathcal{U}_i = \left\{ \sum_{k=1}^{L+1} w_{ik} : \lambda_i \geq 0, \sum_{k=1}^{L+1} \lambda_i^{(k)} = 1, w_{ik}^\top Q_i w_{ik} \leq \Omega_i \left(\lambda_i^{(k)} \right)^2 \forall k \right\}.$$

Existing RO techniques can not handle this convex uncertainty region because of its nonconvex representation, and do not provide ways to reformulate the “for all” constraint. Our method results in a formulation without semi-infinite constraints. Even though it results in a nonconvex formulation, global optimization techniques can be applied to solve the problem.

4. The fourth example of new tractable uncertainty regions is illustrated by the following robust constraint:

$$\mathbf{a}_i^\top \mathbf{x} + \sum_j h_{ij}(a_{ij}) x_j \leq b_i \quad \forall \mathbf{a}_i : \|\mathbf{a}_i\|_\infty \leq 1,$$

where the functions h_{ij} are convex. For many choices of h_{ij} this constraint is not tractable. To show that it can be solved with our method, we first move the nonlinearity to the uncertainty region:

$$\mathbf{a}_i^\top \mathbf{x} + \sum_j d_{ij} x_j \leq b_i \quad \forall (\mathbf{a}_i, \mathbf{d}_i) : \|\mathbf{a}_i\|_\infty \leq 1, \quad h_{ij}(a_{ij}) = d_{ij} \quad \forall j,$$

and then obtain an equivalent constraint by taking the convex hull of the uncertainty region:

$$\begin{aligned} \mathbf{a}_i^\top \mathbf{x} + \sum_j d_{ij} x_j \leq b_i \quad \forall (\mathbf{a}_i, \mathbf{d}_i) : \|\mathbf{a}_i\|_\infty \leq 1, \quad h_{ij}(a_{ij}) \leq d_{ij}, \\ 2d_{ij} \leq h_{ij}(1)(a_{ij} + 1) - h_{ij}(-1)(a_{ij} - 1) \quad \forall j. \end{aligned}$$

This transformation has also been applied in [4, p. 20], but they require a closed form for the convex conjugate of h_{ij} to reformulate this constraint. With our method, this linear constraint with a convex uncertainty region is tractable for any convex h_{ij} .

4 Globalized Robust Counterpart

A robust constraint holds for all realizations of the uncertain parameters in the uncertainty region. This may be very pessimistic. [3] proposes the globalized robust counterpart (GRC), which may be used to reduce the conservatism of the RC (also see [6, Ch. 3]). Let $\mathcal{U}_i = \{\mathbf{a}_i : f_{ik}(\mathbf{a}_i) \leq 0 \quad \forall k\}$ be the set of “physically possible” realizations, and let a smaller set $\mathcal{U}'_i = \{\mathbf{a}_i : g_{ik}(\mathbf{a}_i) \leq 0 \quad \forall k\} \subset \mathcal{U}_i$ contain the “normal range” of realizations. We define the GRC as:

$$\mathbf{a}_i^\top \mathbf{x} \leq b_i + \min_{\mathbf{a}_i' \in \mathcal{U}'_i} \{h_i(\mathbf{a}_i, \mathbf{a}_i')\} \quad \forall \mathbf{a}_i \in \mathcal{U}_i, \quad (7)$$

where h_i is a nonnegative jointly convex function for which $h_i(\mathbf{a}_i', \mathbf{a}_i') = 0$ for all \mathbf{a}_i' in \mathcal{U}_i' . Examples are norms and ϕ -divergence measures. The term after b_i in (7) denotes the allowable violation of the constraint, which is 0 if \mathbf{a}_i is in the smaller set \mathcal{U}_i' . This definition of the GRC is more general than in [6]: the set \mathcal{U}_i does not have to be the Minkowski sum of a convex region and a cone, and the allowable violation does not have to be linear in a norm. Our definition also measures the distance in a more natural way than in [6]. Suppose the uncertainty region is the Minkowski sum of a circle around $(0,0)$ with radius 1 and the nonnegative orthant. In this case, $\mathcal{U}_i = \{\mathbf{a} + \mathbf{b} : \|\mathbf{a}\|_2 \leq 1, \mathbf{b} \in \mathbb{R}_+^n\}$ and $\mathcal{U}_i' = \{\mathbf{a} : \|\mathbf{a}\|_2 \leq 1\}$. The distance between $(-1,1)$ and \mathcal{U}_i' is 1 in the definition of [6], because the difference vector has to be an element from the cone. With our definition, the distance may be the Euclidean distance $\sqrt{2} - 1$.

We will show that (7) can be reformulated to a linear constraint with a convex uncertainty region. Constraint (7) is equivalent to:

$$\mathbf{a}_i^\top \mathbf{x} \leq b_i + d_i \quad \forall (\mathbf{a}_i, d_i) : f_{ik}(\mathbf{a}_i) \leq 0 \forall k \quad d_i = \min_{\mathbf{a}_i' \in \mathcal{U}_i'} \{h_i(\mathbf{a}_i, \mathbf{a}_i')\},$$

which in turn is equivalent to:

$$\mathbf{a}_i^\top \mathbf{x} \leq b_i + d_i \quad \forall (\mathbf{a}_i, \mathbf{a}_i', d_i) : f_{ik}(\mathbf{a}_i) \leq 0 \forall k \quad g_{ik}(\mathbf{a}_i') \leq 0 \forall k \quad d_i \geq h_i(\mathbf{a}_i, \mathbf{a}_i').$$

This is indeed a linear constraint with a convex uncertainty region. We will now show how the GRC can be formulated. Consider the following Globalized Robust program:

$$\begin{aligned} \text{(R-GRC)} \quad & \max_{\mathbf{x}} \quad \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{a}_i^\top \mathbf{x} \leq b_i + d_i \quad \forall (\mathbf{a}_i, \mathbf{a}_i', d_i) : f_{ik}(\mathbf{a}_i) \leq 0 \forall k, \\ & g_{ik}(\mathbf{a}_i') \leq 0 \forall k \quad d_i \geq h_i(\mathbf{a}_i, \mathbf{a}_i') \quad \forall i \\ & \mathbf{x} \in \mathcal{K}, \end{aligned}$$

whose optimistic dual is:

$$\begin{aligned} \text{(OD-GRC)} \quad & \min_{\mathbf{a}_i, \mathbf{a}_i', d_i, \mathbf{y}} \quad \mathbf{b}^\top \mathbf{y} + \mathbf{d}^\top \mathbf{y} \\ \text{s.t.} \quad & \sum_i y_i \mathbf{a}_i - \mathbf{c} \in \mathcal{K}^* \\ & f_{ik}(\mathbf{a}_i) \leq 0 \quad \forall i \quad \forall k \quad (8) \\ & g_{ik}(\mathbf{a}_i') \leq 0 \quad \forall i \quad \forall k \quad (9) \\ & h_i(\mathbf{a}_i, \mathbf{a}_i') \leq d_i \quad \forall i \quad (10) \\ & \mathbf{y} \geq 0. \quad (11) \end{aligned}$$

We substitute $\mathbf{v}_i = y_i \mathbf{a}_i$, $\mathbf{v}_i' = y_i \mathbf{a}_i'$ and $w_i = y_i d_i$ and multiply constraints (8)–(10) with y_i :

$$\begin{aligned} \text{(COD-GRC)} \quad & \min_{\mathbf{v}_i, \mathbf{v}_i', w_i, \mathbf{y}} \quad \mathbf{b}^\top \mathbf{y} + \sum_i w_i \\ \text{s.t.} \quad & \sum_i \mathbf{v}_i - \mathbf{c} \in \mathcal{K}^* \\ & y_i f_{ik}(\frac{\mathbf{v}_i}{y_i}) \leq 0 \quad \forall i \quad \forall k \\ & y_i g_{ik}(\frac{\mathbf{v}_i'}{y_i}) \leq 0 \quad \forall i \quad \forall k \\ & y_i h_i(\frac{\mathbf{v}_i}{y_i}, \frac{\mathbf{v}_i'}{y_i}) \leq w_i \quad \forall i \quad (12) \\ & \mathbf{y} \geq \mathbf{0}. \end{aligned}$$

Note that the product $y_i \mathbf{a}_i'$ does not appear in (OD-GRC), but that the substitution $\mathbf{v}_i' = y_i \mathbf{a}_i'$ is still necessary to make (COD-GRC) convex. For the tractability of (COD-GRC), all results regarding the functions that define the uncertainty region \mathcal{U}_i also apply to the functions that define \mathcal{U}_i' . When h_i is a ϕ -divergence measure, constraint (10) is given by:

$$\sum_j a_{ij} \phi(a'_{ij}/a_{ij}) \leq d_i.$$

Constraint (12) then contains the perspective of the perspective of ϕ , which is just the perspective:

$$\sum_j v_{ij} \phi(v'_{ij}/v_{ij}) \leq w_i.$$

When h_i is an arbitrary norm $\|\mathbf{a}_i - \mathbf{a}_i'\|$, constraint (12) contains the same norm: $\|\mathbf{v}_i - \mathbf{v}_i'\| \leq w_i$.

5 Multi-item newsvendor example

We demonstrate our new method on a robust LP with a convex uncertainty region that currently can not be solved with other methods. We slightly modify the multi-item newsvendor problem described in [5], because without modification we can not show the substitution. There are 12 items indexed by i , and each has its own ordering cost c_i , selling price v_i , salvage price r_i , and unsatisfied demand loss l_i . So, when the order quantity Q_i is less than the demand d_i , the profit equals $v_i Q_i + l_i(Q_i - d_i) - c_i Q_i$, and $v_i d_i + r_i(Q_i - d_i) - c_i Q_i$ otherwise. If $r_i \leq v_i + l_i$, this profit is concave piecewise linear in the decision variable Q_i . In practice the demand is not known, but for every item we can define scenarios s which occur for item i which will occur independently of other items with probability $p_i^{(s)}$, resulting in a demand of $d_i^{(s)}$. The goal is to determine Q_i such that the total ordering cost is minimized while having an expected profit of at least γ . This can be formulated as a robust LP as follows:

$$\begin{aligned} \text{(R-NV)} \quad & \min_{\mathbf{Q}, \mathbf{u}} \quad \sum_i c_i Q_i \\ \text{s.t.} \quad & \sum_i \sum_s p_i^{(s)} u_i^{(s)} \geq \gamma \\ & \forall \mathbf{p}_i \geq \mathbf{0} : \sum_s p_i^{(s)} = 1, \sum_s \left| \left(\hat{p}_i^{(s)} \right)^\alpha - \left(p_i^{(s)} \right)^\alpha \right|^{1/\alpha} \leq \rho, \quad \forall i \quad (13) \\ & u_i^{(s)} + (c_i - r_i) Q_i \leq d_i^{(s)} (v_i - r_i) \quad \forall i \quad \forall s \\ & u_i^{(s)} + (c_i - v_i - l_i) Q_i \leq -d_i^{(s)} l_i \quad \forall i \quad \forall s \\ & \mathbf{Q} \geq \mathbf{0}, \end{aligned}$$

where $u_i^{(s)}$ denotes the profit for item i in scenario s , and the uncertainty region is based on the Matusita distance with α in $(0, 1)$. The optimistic dual of (R-NV) is

given by:

$$\begin{aligned}
(\text{OD-NV}) \quad & \max_{x, \mathbf{y}, \mathbf{z}, p} \quad \gamma x + \sum_{i,s} d_i^{(s)} ((v_i - r_i) y_{is} - l_i z_{is}) \\
\text{s.t.} \quad & p_i^{(s)} x + y_{is} + z_{is} = 0 \quad \forall i \quad \forall s \\
& \sum_s \{(c_i - r_i) y_{is} + (c_i - v_i - l_i) z_{is}\} \leq c_i \quad \forall i \\
& \sum_s p_i^{(s)} = 1 \quad \forall i
\end{aligned} \tag{14}$$

$$\begin{aligned}
& \sum_s \left| \left(\hat{p}_i^{(s)} \right)^\alpha - \left(p_i^{(s)} \right)^\alpha \right|^{1/\alpha} \leq \rho \quad \forall i \\
& x \geq 0, \quad \mathbf{y} \leq \mathbf{0}, \quad \mathbf{z} \leq \mathbf{0}, \quad p \geq 0.
\end{aligned} \tag{15}$$

After substituting $w_{is} = p_i^{(s)} x$ and multiplying (14) and (15) with x , the convex reformulation becomes:

$$\begin{aligned}
(\text{COD-NV}) \quad & \max_{x, \mathbf{y}, \mathbf{z}, w} \quad \gamma x + \sum_{i,s} d_i^{(s)} ((v_i - r_i) y_{is} - l_i z_{is}) \\
\text{s.t.} \quad & w_{is} + y_{is} + z_{is} = 0 \quad \forall i \quad \forall s
\end{aligned} \tag{16}$$

$$\sum_s \{(c_i - r_i) y_{is} + (c_i - v_i - l_i) z_{is}\} \leq c_i \quad \forall i \tag{17}$$

$$\sum_s w_{is} = x \quad \forall i$$

$$\sum_s \left| \left(\hat{p}_i^{(s)} x \right)^\alpha - w_{is}^\alpha \right|^{1/\alpha} \leq \rho x \quad \forall i \tag{18}$$

$$x \geq 0, \quad \mathbf{y} \leq \mathbf{0}, \quad \mathbf{z} \leq \mathbf{0}, \quad w \geq 0.$$

Note that $x = 0$ implies $w_{is} = 0$, which is in accordance with the substitution.

We take $\alpha = 0.5$, $\gamma = 100$ and for all other parameters we take the same values as reported in [5]. This means that there are three scenarios, corresponding to low ($d_i^{(s)} = 4$), medium ($d_i^{(s)} = 8$) and high ($d_i^{(s)} = 10$) demand. The other parameters are listed in Table 1. We solve the problem for different values of ρ , varying between 0.000 and 0.030 in steps of 0.0001, with AIMMS 3.11 and KNITRO 7.0. For values of ρ larger than 0.0306, the problem is infeasible. The robust optimal \mathbf{Q} and \mathbf{u} are the elements of a KKT vector corresponding to constraints (16) and (17), respectively. The optimal order quantities and corresponding ordering costs are listed in Table 2 for different values of ρ . For a given solution, the objective values of (R-NV) and (COD-NV) should be equal. We observe both positive and negative differences of at most 0.02%, probably due to numerical limitations of the solver. The constraint violation of (13) can be computed by maximizing a linear function over a convex set, and was found to be at most $1.5 \cdot 10^{-5}$ among all solutions.

For every solution we have uniformly sampled 10,000 p matrices from the uncertainty region, and computed the corresponding expected profit. Because implementing the robust solution requires a larger investment, the following comparison is based on the expected return, which is obtained by dividing the expected profit by the total ordering costs. The mean value and the range of these expected returns are listed in Table 3. As can be seen from this picture, the mean value for the nonrobust solution is often worse than the worst case for the robust solution. We will explain why the robust solution performs much better using the expected return of a single item. In the same way as for all items together, we have computed the expected return for item 3 (Figure 1). The largest increase in expected return is between $\rho = 0$ and $\rho = 0.005$, for which the order quantity increases from 4.00 to 6.20 (Table 2). The profits for item 3 in the three scenarios are (12, -8, -18) for $Q_3 = 4.00$, and (2.1, 9.6, -0.4) for $Q_3 = 6.20$. So,

with a slightly larger investment, the variation of the profit becomes much smaller, and hence, deviations in the probabilities on the scenarios have a smaller impact on the expected profit. This reduces the range of the expected return of the the robust solution, which can be seen in Figure 1.

Table 1: Parameter values for the multi-item newsvendor example

i	1	2	3	4	5	6	7	8	9	10	11	12
c_i	4	5	6	4	5	6	4	5	6	4	5	6
v_i	6	8	9	5	9	8	6	8	9	6.5	7	8
r_i	2	2.5	1.5	1.5	2.5	2	2.5	1.5	2	2	1.5	1
l_i	4	3	5	4	3.5	4.5	3.5	3	5	3.5	3	5
$\hat{p}_i^{(1)}$	0.375	0.250	0.375	0.127	0.958	0.158	0.485	0.142	0.679	0.392	0.171	0.046
$\hat{p}_i^{(2)}$	0.375	0.250	0.250	0.786	0.007	0.813	0.472	0.658	0.079	0.351	0.484	0.231
$\hat{p}_i^{(3)}$	0.250	0.500	0.375	0.087	0.035	0.029	0.043	0.200	0.242	0.257	0.345	0.723

Table 2: Optimal ordering cost and quantities for the multi-item newsvendor problem

ρ	cost	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}	Q_{11}	Q_{12}
0.000	391	8.00	8.00	4.00	8.00	4.00	8.00	4.00	8.00	4.00	8.00	7.03	8.00
0.005	412	8.00	8.00	5.87	8.00	4.00	8.00	5.69	8.00	4.00	7.01	8.00	8.34
0.010	421	8.00	8.00	6.20	8.00	4.00	8.00	6.12	8.00	4.00	7.55	8.00	8.85
0.015	430	8.00	8.00	6.39	8.00	4.00	8.00	6.36	8.00	4.00	8.00	8.00	9.62
0.020	440	8.00	8.00	7.10	8.00	4.00	8.00	7.31	8.00	4.00	8.00	8.00	10.00
0.025	453	8.00	8.00	7.36	8.00	4.00	8.00	8.00	8.00	5.51	8.00	8.00	10.00
0.030	469	8.00	9.49	8.00	8.00	4.00	8.00	8.00	8.00	6.26	8.00	8.00	10.00

Table 3: Simulation results of the expected return for the robust and nonrobust solutions for the multi-item newsvendor problem.

ρ	Robust solution			Nonrobust solution		
	min	mean	max	min	mean	max
0.000	0.2557	0.2557	0.2557	0.2557	0.2557	0.2557
0.005	0.2538	0.2695	0.2842	0.2369	0.2542	0.2737
0.010	0.2545	0.2753	0.2987	0.2269	0.2529	0.2795
0.015	0.2560	0.2806	0.3076	0.2169	0.2515	0.2855
0.020	0.2502	0.2855	0.3215	0.2121	0.2501	0.2903
0.025	0.2497	0.2815	0.3125	0.2088	0.2490	0.2955
0.030	0.2425	0.2809	0.3140	0.2026	0.2476	0.2900

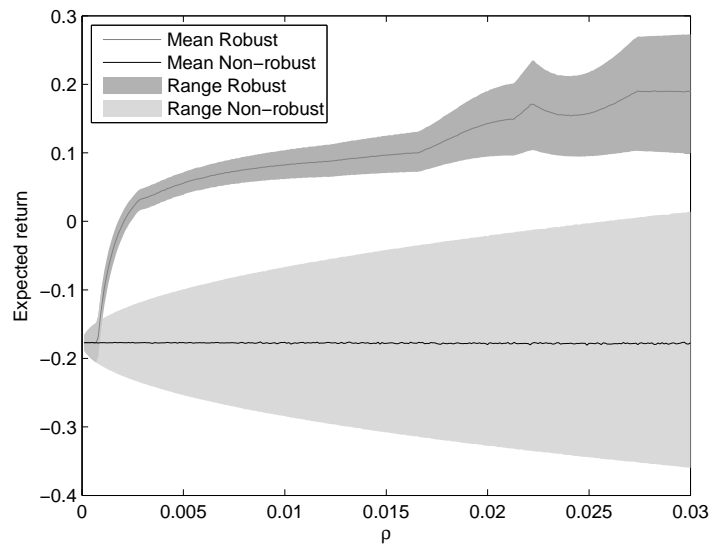


Figure 1: Simulation results of the expected return for item 3 in the robust and nonrobust solutions for the multi-item newsvendor problem.

References

- [1] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh. Clustering with Bregman divergences. *The J. of Mach. Learn. Res.*, 6:1705–1749, 2005.
- [2] A. Beck and A. Ben-Tal. Duality in robust optimization: primal worst equals dual best. *Oper. Res. Lett.*, 37(1):1–6, 2009.
- [3] A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems. *Math. Prog.*, 107(1):63–89, 2006.
- [4] A. Ben-Tal, D. den Hertog, and J.-Ph. Vial. Deriving robust counterparts of nonlinear uncertain inequalities. *CentER Discuss. Pap.*, 2012(053), 2012.
- [5] A. Ben-Tal, D. den Hertog, A. M. B. De Waegenaere, B. Melenberg, and G. Rennen. Robust solutions of optimization problems affected by uncertain probabilities. *CentER Discuss. Pap.*, 2011(061), 2011. To appear in *Manag. Sci.*
- [6] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton Ser. in Appl. Math. Princeton University Press, Princeton, NJ, 2009.
- [7] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Oper. Res. Lett.*, 25(1):1–13, 1999.
- [8] D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM Rev.*, 53(3):464–501, 2011.
- [9] H. Blanc and D. den Hertog. On Markov chains with uncertain data. *CentER Discuss. Pap.*, 2008(050), 2008.
- [10] L. M. Bregman. The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. and Math. Phys.*, 7(3):200–217, 1967.
- [11] B. Dacorogna and P. Maréchal. The role of perspective functions in convexity, polyconvexity, rank-one convexity and separate convexity. *J. Convex Anal.*, 15(2):271–284, 2008.
- [12] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ, 1963.
- [13] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM J. on Matrix Anal. and Appl.*, 18(4):1035–1064, 1997.
- [14] J. E. Falk. Exact solutions of inexact linear programs. *Oper. Res.*, 24(4):783–787, 1976.
- [15] R. N. Kaul, S. Kaur, and V. Lyall. Duality in inexact fractional programming with set-inclusive constraints. *J. of Optim. Theory and Appl.*, 50(2):279–288, 1986.
- [16] K. Matusita. On the notion of affinity of several distributions and some of its applications. *Ann. of The Inst. of Stat. Math.*, 19(1):181–192, 1967.
- [17] A. Rényi. On measures of entropy and information. In *Proc. of the 4th Berkeley Symp. on Math., Stat. and Probab.*, volume I, pages 547–561. Univ. of California Press, 1961.
- [18] R. T. Rockafellar. *Convex Analysis*. Princeton Univ. Press, Princeton, NJ, 1970.

- [19] M. Römer. Von der Komplexmethode zur robusten Optimierung und zurück. In *Beiträge zum Festkolloquium "Angewandte Optimierung" anlässlich des 65. Geburtstages von Prof. Dr. Rolf Rogge*, pages 37–61. Universität Halle–Wittenberg, 2010.
- [20] C. Singh. Convex programming with set-inclusive constraints and its applications to generalized linear and fractional programming. *J. of Optim. Theory and Appl.*, 38(1):33–42, 1982.
- [21] A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Oper. Res.*, 21(5):1154–1157, 1973.
- [22] A. L. Soyster. A duality theory for convex programming with set-inclusive constraints. *Oper. Res.*, 22(4):892–898, 1974.
- [23] D. J. Thuente. Duality theory for generalized linear programs with computational methods. *Oper. Res.*, 28(4):1005–1011, 1980.